How to use machine learning to discover relevant aspects in the Dengue Virus protease and guide the drug discovery process?

Dengue Virus (DENV) is perhaps one of the most relevant viral pathogens in tropical and sub-tropical areas. A World Health Organization (WHO) report indicates that DENV related infections have increased 30-fold in the last 50 years. DENV circulates as four different serological groups, between 50 and 100 million people are affected by DENV every year, and 2.5 billion people live in an area of risk. The many people at risk is highly relevant to assess the impact of the virus since some DENV related infections as Dengue hemorrhagic fever and Dengue Shock Syndrome could have a mortality rate as high as 2.5%.

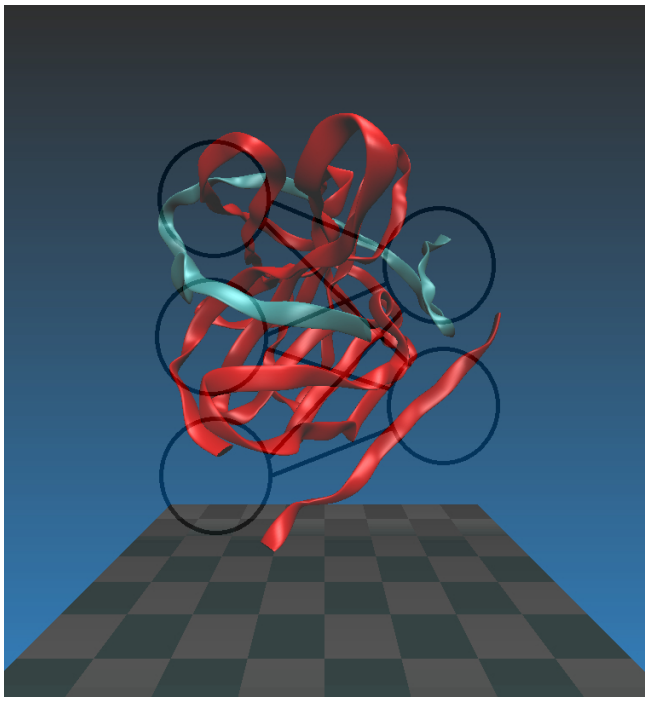

Fig. 1. In red, Dengue virus NS3 protease, and in cyan its NS2b cofactor. Machine learning methods were used to characterize novel sites in the protein complex as susceptible targets or anchor points for drugs against all four serotypes of Dengue virus.

Since November 2015 there is a vaccine available to prevent DENV infection. However, there is currently no antiviral drug available for DENV infection treatment. Many experts have indicated that an antiviral drug could be helpful to manage infections in vulnerable groups like children and the elderly. Also, it could be beneficial to slow down DENV spreading and treatment of severe cases. In the past decade, there have been many efforts to discover new DENV inhibitors with little

success partly because the known targets that are being used to guide the antiviral drug development face different challenges to the discovery process.

Dengue virus belongs the Flavivirus genus, same as Zika Virus and West Nile Virus. As part of its life-cycle, it produces a single protein (polyprotein) that after cleavage creates all the viral proteins needed to accomplish its function inside the infected cell and to assemble the new viruses. Cleavage is essential for the virus to generate functional proteins. The Viral Ns2b-NS3 Protease complex and host cell enzymes oversee cleaving the polyprotein. Therefore, DENV Ns2b-NS3 Protease is an excellent candidate to guide the drug discovery process.

In structure-based drug design, a known target is required to guide the discovery process. However, it is not enough to know the 3D structure of the target, in this case, the protein. It is essential to know the specific residues that have a functional relevance to discover a molecule that can bind them. In the case of the Ns2b-NS3 Protease, only a few functional relevant residues were known, and they were all related directly to the active site and therefore the catalytic role of the enzyme. Many authors have stated that because of the active site nature, it is hard to develop an adequate drug. Within this scenario, our team decided to use computational methods to discover new functionally relevant residues in the DENV Ns2b-NS3 Protease that could guide the drug discovery process.

In this case, we described a step by step process to build a machine learning algorithm capable of identifying new functionally relevant residues in the DENV Ns2b-NS3 Protease. We fed our algorithm with published experimental data available and different residue-based molecular descriptors (computationally calculated). It is worth noting that some of these molecular descriptors were obtained with other publicly accessible Machine Learning methods. Moreover, the methodology described in this work could virtually be applied to any biological system. Therefore, we identified some residues that clustered together outside the active site of the protease. These residues could be useful to guide the development of noncompetitive inhibitors for the Ns2b-NS3 protease and therefore a potential drug to treat DENV infections. Finding functionally relevant sites outside the active site is relevant since DENV nonstructural proteins have different non-canonical roles different than their catalytic functions directly related to the viral cycle.

Daniel Aguilera-Pesantes, Miguel A. Mendez
Universidad San Francisco de Quito, Instituto de Simulacion Computacional, Ecuador

Publication

Discovering key residues of dengue virus NS2b-NS3-protease: New binding sites for antiviral inhibitors design.
Aguilera-Pesantes D, Robayo LE, Méndez PE, Mollocana D, Marrero-Ponce Y, Torres FJ, Méndez

M
Biochem Biophys Res Commun. 2017 Oct 28